ED 174 458 SE 028 528

AUTHOR Treasure, Morris
TITLE Alberta Science Achievement Study: A Study Conducted for the Minister's Advisory Committee on Student Achievement. Executive Summary.
INSTITUTION Alberta Dept. of Education, Edmonton.
PUB DATE 78
NOTE 17p.; Contains occasional light type

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *Data Bases; Educational Assessment; *Educational Research; *Elementary School Science; Elementary School Students; Elementary Secondary Education; Evaluation; *Science Education; Sciences; Science Tests; *Secondary School Science; Secondary School Students

ABSTRACT

This report is an executive summary of a study conducted for the Minister's Advisory Committee on Student Achievement (MACOSA) of Alberta, Canada. It was designed to investigate levels of student achievement in science in Alberta at grades three, six, nine and twelve and to provide a data base for future assessment. Between 2,000 and 3,000 students were tested at each grade level. The findings show that the student performances were generally satisfactory. Some areas of weakness were identified: (1) knowledge of scientific methods in grade three; (2) knowledge of physical science in grades three, six and twelve; and (3) earth-space science and general knowledge of science and scientists in grades nine and twleve. A list of recommendations is presented at the end of the report. (HM)

Alberta Science Achievement Study

Executive Summary

MARCH, 1979

Alberta

Minister's Advisory Committee on Student Achievement

ALBERTA SCIENCE ACHIEVEMENT STUDY

A STUDY CONDUCTED FOR
THE MINISTER'S ADVISORY COMMITTEE ON
STUDENT ACHIEVEMENT

Executive Summary

- by -

Morris Treasure
Planning and Research Branch
Alberta Education

Fall, 1978

To speculate without facts is to attempt to enter a house of which one has not the key, by wandering aimlessly round and round, searching the walls, now and then peeping through the windows. Facts are the key.

- Julian Huxley
Essays in Popular Science

The Minister's Advisory Committee on Student Achievement (MACOSA) was established by ministerial order in October 1976 in response to growing concerns expressed by the public-at-large, government, labor, business, students and educators regarding the quality and standards of basic education in Alberta.

MACOSA commissioned a number of studies, primarily to provide basic information for a summary of current levels of achievement in Alberta and to provide baseline data for future assessment. These studies fell into three categories: (1) preliminary studies, (2) achievement studies, and (3) other studies.

This achievement study, Alberta Science Achievement Study, was designed to provide information about current levels of achievement in science among students in Alberta schools and to provide a data base for future assessments.

This report, which represents the findings and conclusions of the researcher, was presented to MACOSA as information.

ACKNOWLEDGEMENTS

The advice and assistance of a wide group of colleagues in Alberta Education and the university community was a valuable prerequisite of the study.

The assistance of Dr. J. E. Reid and the staff of the Student Evaluation and Data Processing Branch is publicly acknowledged. The support and assistance of the Planning and Research Branch is also greatly appreciated.

Advice, sympathy and assistance far beyond the usual responsibilities of a departmental committee were received from the following members of the steering committee:

Dr. L. R. Tolman, Chairman
Dr. H. C. Rhodes, Recorder
Dr. W. G. Holliday
Mr. M. Lynch
Mr. I. Ibuki
Dr. H. G. Sherk, Executive Secretary of MACOSA

The cooperation of the participants in the study is also greatly appreciated--the students who took the time to write the tests, the teachers who administered the tests, and those who cared enough to respond in writing. The investigator also acknowledges the assistance of the support staff of the Planning and Research Branch who have typed thousands of pages of tests, charts and reports.

Science Achievement Study

Purposes

The purposes of this study, entitled the Alberta Science Achievement Study, were:

1. To investigate current levels of student achievement in science in Alberta at the grades three, six, nine and twelve levels.

2. To provide a data base for future assessments.

Procedures

For the grade three test the researcher chose 114 items from a variety of sources, and a randomly selected group of 20 primary teachers validated the items. For the grade six test the researcher chose 144 items from similar sources, and a panel of 18 upper elementary teachers validated them. A determined effort was made to select items which would reflect the objectives of the elementary science program. There are six content areas in this program--two in physical science, three in biological science and one in earth science. Interwoven with these content areas are objectives related to the methods of science, and attitudes toward and knowledge of science and scientists.

The tests for grades nine and twelve were made up of two published standardized tests: the Sequential Tests of Educational Progress (STEP), Series II, and Test of Understanding Science (TOUS), Form Jw. The STEP II measures student achievement levels and the TOUS measures students' knowledge of and opinions about scientists and science as a field of study. For grade nine the researcher chose the STEP II, Form 3A (with the addition of 15 supplementary items) because it best matched the Alberta curriculum. In grade twelve, where the emphasis was to be on more general knowledge of science, both in terms of the content of the high school program and the general aims and purposes of science, the researcher chose the STEP II, Form 2A, because of its relative quality and suitability for gathering the required information. The TOUS, Form Jw was the only appropriate attitude test among the few such tests which are available.

The Student Evaluation and Data Processing Branch of Alberta Education selected a stratified random sample of schools consisting of 101 schools offering grade three; 96 schools offering grade six; 40 schools offering grade nine; and 24 schools offering grade twelve. There were 3073 grade three students 2935 grade six students, 2426 grade nine students, and 2125 grade twelve students tested. These represent about 8 percent of the total number of students enrolled in these grades.

All tests were administered on May 17, 1978.

Findings

Table 1 shows the levels of performance for each grade level and general content area, reported as percentage of correct responses, and the number of items used to test each area.

TABLE 1

Summary of Science Achievement Levels: Percentage of Correct Responses

Content Area	Grade 3	Grade 6	Grade 9	Grade 12
Physical Science	55.2 (21)*	59.6 (38)	62.6 (17)	55.5 (32)
Biological and Life Science	63.7 (26)	61.8 (45)	69.3 (12)	72.7 (16)
Earth Science	77.2 (6)	70.8 (11)	56.7 (17)	47.9 (3)
Methods of Science	57.4 (41)	64.9 (24)	62.4 (19)	64.8 (36)
Average Student Achievement in Content Areas	60.3	62.5	62.2	62.2
Science as a Human Endeavor	60.6 (14)	71.1 (14)	49.2 (45)	57.3 (45)
Interest in Science	66.1 (6)	61.2 (12)	----	----

* Number of test items is shown in parentheses.

In grade three the performance levels were highest on earth science, even though the curriculum does not stress earth science; and lowest on physical science, which includes a number of difficult concepts such as electricity, molecules and basic energy conservation. Grade three students registered a relatively low level of achievement in methods of science investigation despite the fact that science programs have emphasized this area over the past few years. These students demonstrated a reasonable amount of knowledge in the area of science as a human endeavor, a section of the test designed to check students' perceptions of science and scientists.

The grade six curriculum has recently emphasized physical science but, even so, grade six scores in physical science were not high. Scores were high on biological science and life science items and as in grade three, students at the grade six level had high scores on the earth science items. Relatively high scores on science methods items probably reflect an increased curricular emphasis on this topic. In grade six scores on perceptions of science and scientists were also relatively high.

Grade nine students achieved the highest average scores on items dealing with biological and life science, and the next highest average score on physical science. The achievement level in earth science was quite low. Scores were high on methods of science while "science as a human endeavor", which included such topics as aims of science and the skills and aptitudes of scientists, had a low response rate.

Grade twelve students achieved at the highest level on biological and life science items and lowest on earth science items. Performances were relatively high on methods of science and relatively low on science as a field of human endeavor.

The tests for grades three and six contained 47 common items, including 18 items drawn from the content areas for grade three, 18 items drawn from grade six content areas and a further 11 items measuring interest in science and opinions and beliefs about scientists. Table 2 indicates relative student performance on common items.

TABLE 2

Performance on Common Items on Science Tests:
Percentage of Correct Responses

	Student Performance (%) Grade 3	Grade 6
Grade 3 Target Items (18)	63.0	77.0
Grade 6 Target Items (18)	49.4	59.1
Interest and Opinion Items (11)	61.4	74.0
TOTAL	57.4	69.4

As expected, grade six student performance levels on the common items were substantially higher than grade three performance levels. However, the grade six performance on grade six items was somewhat low.

Relative performance on common items by content area, as shown in Table 3, is not as consistent.

TABLE 3

Performance on Common Items
on Science Tests by Content Area:
Percentage of Correct Responses

Content	Student Performance (%)	
	Grade 3	Grade 6
Physical Science (5)*	62.9	64.3
Biological and Life Science (10)	58.2	74.2
Earth Science (4)	73.1	86.9
Methods of Science (13)	48.4	58.1
Science as a Human Endeavor (10)	57.5	74.9
Interest in Science (5)	61.1	70.0

* Number of test items is shown in parentheses.

The most substantial differences between performances by grade three and grade six on common items occurred on item clusters dealing with earth science, biological and life science, and science as a human endeavor (opinions and beliefs about scientists), while smaller differences occurred on item clusters dealing with physical science, methods of science and interest in science.

Because both grades nine and twelve students responded to the TOUS, Form Jw, and to different forms of the STEP II, the researcher was able to make similar comparisons between these grades. Table 4 compares student performances in grade nine and twelve on the STEP II, for Alberta and the United States.

TABLE 4

STEP II Test Results:
Average Raw Scores

	Form 3A (Grade 9) Alberta (Spring '78)	Form 3A (Grade 9) U.S.A. (Spring '70)	Form 2A (Grade 12) Alberta (Spring '78)	Form 2A (Grade 12) U.S.A (Spring'70)
N	2426	2637	2125	2285
Average Score	34 (50)	32 (50)	47 (75)	42 (75)
Standard Deviation	7	12	12	13

* Total number of test items is shown in parentheses.

The average response of grade nine students on the STEP II was 34 out of 50 items, as compared with 32 for their American counterparts. The average score for the grade twelve students was 47 out of 75 items, as compared with 42 for the American norming sample. At both grade levels the high Alberta performances are statistically significant ($p < 0.05$).

Table 5 compares performances by students in grades nine and twelve with the results further broken down to show performance levels for girls and boys.

TABLE 5

TOUS, Form Jw, Test Results:
Average Raw Scores

	Grade 9	Boys	Girls	Grade 12	Boys	Girls
N	2426	1234	1144	2125	1059	1034
Average Score	21.8 (45)*	21.1	22.6	25.3 (45)*	24.6	25.8
Standard Deviation	6.0	6.1	5.8	6.5	6.7	6.2

* Total number of test items is shown in parentheses.

The TOUS has not been used in Alberta long enough to make comparisons other than those shown in Table 5. As one might expect, grade twelve students scored higher than grade nine students. In both grades nine and twelve, the girls outperformed the boys. This result suggests that girls understand the nature of science better than do boys.

An examination of grades nine and twelve results by content area (not shown in table form) indicated that both groups performed at a high level on the biological and life science items and at a fairly low level on items asking about the nature of science (interest and opinion items). Both grades performed above the test average on items related to scientific methods, and average performance on physical science items was above the median for grade nine but below the median for grade twelve.

The researcher also categorized the test items by the thought level required for giving the correct response. Table 6 shows the average performance for three thought levels--knowledge, comprehension and application.

TABLE 6

Average Performance by Thought Level of Items:
Percentage of Correct Responses

Grade	Knowledge	Thought Level Comprehension	Application
3	68.3 (16)*	57.8 (24)	67.0 (15)
6	67.4 (38)	57.7 (32)	59.3 (26)
9	64.7 (13)	69.6 (18)	58.6 (25)
12	61.5 (15)	63.1 (15)	58.6 (20)

* Number of items is shown in parentheses.

Elementary students had higher scores than secondary students at the knowledge level, but lower scores than secondary students on comprehension items. Scores on items at the application level were substantially lower than scores at the knowledge level for all grades except grade three.

Conclusions

These data show the present level of student achievement in Alberta science programs on a number of dimensions.

An examination of grades nine and twelve results by content area (not shown in table form) indicated that both groups performed at a high level on the biological and life science items and at a fairly low level on items asking about the nature of science (interest and opinion items). Both grades performed above the test average on items related to scientific methods, and average performance on physical science items was above the median for grade nine but below the median for grade twelve.

The researcher also categorized the test items by the thought level required for giving the correct response. Table 6 shows the average performance for three thought levels--knowledge, comprehension and application.

TABLE 6

Average Performance by Thought Level of Items:
Percentage of Correct Responses

Grade	Knowledge	Thought Level Comprehension	Application
3	68.3 (16)*	57.8 (24)	67.0 (15)
6	67.4 (38)	57.7 (32)	59.3 (26)
9	64.7 (13)	69.6 (18)	58.6 (25)
12	61.5 (15)	63.1 (15)	58.6 (20)

* Number of items is shown in parentheses.

Elementary students had higher scores than secondary students at the knowledge level, but lower scores than secondary students on comprehension items. Scores on items at the application level were substantially lower than scores at the knowledge level for all grades except grade three.

Conclusions

These data show the present level of student achievement in Alberta science programs on a number of dimensions.

Whether this level is good, bad or average is very difficult to judge. The descriptive information is provided by this study to provide a basis for comparison, base-line data points for some future assessment. The value of this study will be determined at that future date.

The lower performance of the grade three students on the physical science items was thought to be acceptable because it has been recognized that both teachers and students have experienced some difficulty with this area of the curriculum. The strong performances in both life science and earth-space science were judged to be very satisfying. It could be hypothesized that these student performances were influenced by the relative emphasis on the space theme by children's television programs and the efforts of such organizations as National Geographic in producing specials on the life science theme.

The only area in elementary science thought to be somewhat unsatisfactory was the student performance on items related to science methods. This poorer performance level is probably due to a different view of the role of student activity in science. Otherwise, student performance was generally satisfactory in the content dimensions of the program. However, the secondary student performance on the items related to the earth-space sciences was judged to be somewhat inadequate. One possible reason for this relatively poor student performance could be the level of abstraction of the concepts.

There is some concern about the adequacy of student performance on items related to the methods of scientists. Student performance levels at both grades nine and twelve were above the test average, so these were viewed as satisfactory. The program objectives related to this area are often neither accepted nor understood as being important. Another area of weak student performance was in understanding the nature of science, as tested by the Test of Understanding Science. This area is difficult to teach and is often viewed as being peripheral to the main intent of science programs.

In general terms, the student performances in grades three and six were satisfactory. Also in general terms, the student performances at grades nine and twelve were satisfactory with the exception of the junior high earth science area.

A further measure of adequacy was available at the secondary level because of the use of a standardized test with U.S. norms. The average student performance in Alberta at grades nine and twelve was above that of the 1970 U.S. norms. Since 1970 there has been a documented decline in standardized test scores across the U.S., so this difference suggests a much greater difference in performance could be shown if more recent norms were available. Using this as a standard, student performance in the field of secondary science in Alberta is satisfactory.

The investigation therefore concluded that, on balance, student achievement is satisfactory. A number of weak areas have been identified, and there are a number of areas in which student performance is quite strong.

The areas of some weakness include physical science in grades three and six, earth science in grade nine, and the grades nine the twelve responses to items asking about the purposes and aims of science as a field of human activity.

Areas in which there is some degree of satisfaction are those dealing with the methods and strategies of science and the life sciences.

1.10 Recommendation

The investigator recommends:

1.10.1 that assessment of the science program in Alberta be a continuing process using the appropriate data from the present study as a base-line against which future achievement levels can be evaluated.

The main value of undertaking an assessment of student achievement at this time lies in the use of the data as a point of comparison for future assessments. The present assessment efforts should represent a beginning point in the on-going evaluation of the education program in Alberta schools.

1.10.2 that computerized item banking be developed and maintained.

It is recognized that the instrumentation developed and/or purchased for this assessment was not wholly satisfactory and that further work on many of the items needed to make them fully acceptable measures of achievement. The establishment of a computerized item bank would facilitate this process by simplifying access to the items and the accumulation of data about the items.

When an acceptable bank of items categorized according to program objectives is available, tests may easily be constructed to measure student achievement along a number of program dimensions. This data can then be accumulated to compile annual or biennial reports. By making the bank available to schools and teachers, the quality of testing in the province can be improved. The teacher gains access to a bank of proven items, and the banked items can be improved and extended by using input from teachers to develop and improve them.

Information about student performance on specific curricular objectives could be made available to curriculum developers on short notice. There could be the capability to respond very quickly to requests for information from those with a legitimate need.

1.10.3 that a computerized bank of items used for the grades three and six testing be extended and improved.

The items used at the elementary level were proven items from other assessment programs which measured achievement of a few selected objectives. Coverage of the program objectives should be broadened by the inclusion of more items. In general, the items were technically adequate but revision of some items could make them more applicable to the Alberta program.

1.10.4 that computerized bank of items be developed for grades nine and twelve.

To expedite the administration of the testing program, a commercially developed test was used to collect the base-line data. This provided some valuable insights into levels of student achievment in Alberta. But as with any instrument developed to sample the broad domain of program objectives, there are a few problems with making inferences about the quality of student achievement in Alberta. If an item bank were developed for Alberta, a better match between the test and the curriculum could be achieved.

A start on a bank of Alberta-valid items could be made by including items from former grade nine departmentals, other assessment programs, and locally-developed tests.

The grade twelve item bank should focus less on the content dimension and more on the practical level of scientific and technical knowledge and applications that should be expected of all graduates of the Alberta education system.

1.10.5 that the scope of the items in an item bank should continue to include the full spectrum of program objectives, including those in the affective domain.

The exclusion of any particular program objectives may well infer that a lesser importance is placed on such objectives. It would be unfortunate if the weighting of curricular objectives came as a result of an unplanned and unconscious distribution of items. Student attitude towards science is another objective that should not be neglected.

1.10.6 that standard~~iz~~ed tests of educational progress should continue to be administered in order to make comparisons beyond provincial boundaries.

It is too easy to become parochial and narrowly focussed in our view of program. To counter this tendency, a regular sampling of student achievement should be instituted as part of the continuing process of program evaluation.

1.10.7 that the provincial testing program be conducted at a time of year which is most convenient.

Separation of the provincial assessment program from school-based student evaluation procedures would serve to lessen the impact of over-testing in the year-end period. In situations where it is reasonable to return test results to the teacher, it would be at a time when the teacher can use such results to modify instruction.

1.10.8 that an item validation procedure be instituted to capitalize on the lessons learned from the Results Interpretation Panels.

The use of a broadly based reaction panel for the purpose of validating the content of the test items for use in either a bank or a provincially administered test is recommended. Further involvement of a community-based group of individuals would serve to widen the scope of provincial evaluations beyond the rather narrow view held by some curriculum specialists. If the aim of schooling is to produce an "educated" student, in the broader sense of community expectation, then representatives of that community should be involved in a reaction role at the initial steps of the evaluation and not just at the final stage.

1.10.9 that a longitudinal study of a particular group of students (cohort) along a few broad curricular goals such as the scientific process be undertaken to investigate the grade placement of the specific objectives and their match with student competencies.

Much of present science program has evolved as a result of experiences elsewhere and the availability of certain publishers' materials. It would be both logical and beneficial to gather information about the present nature of our science program before we expend resources on developing program materials or strategies to change the science program. The information gleaned from such a study would be of great benefit to the revision committees.

1.10.10 that a further analyses of the data collected in this assessment be undertaken by qualified researchers either within Alberta Education or in the educational community.

The anonymity of students and schools in the study would be fully protected, but there are many cross-correlations and factor analyses which were not made and which could in fact lead to an improved assessment.

1.10.11 that a study of the effect of variables such as size of school, presence (or absence) of laboratories, and amount of time spent on science be commissioned by Alberta Education.

The present study was limited in its scope by the very nature of the questions to be answered. However, this kind of process-product information is of value in examining some of the reasons for differences in student performance.

1.10.12 that the testing procedures using matrix-sampling be continued in future province-wide assessments of student performance.

The matrix-sampling procedure proved to be a very efficient, economical way of gathering valid data from across the province on a broad spectrum of objectives. One has to recognize that although a provincial testing program has little attraction for either teachers or students, the cooperation of these two groups is vital to its success. Therefore, limiting student time and minimizing the effort required of the teacher pays off in a greater degree of cooperation. From a provincial perspective, matrix sampling was facilitative. However, from a local perspective (if feedback is to be provided locally about individual pupils) it might not be applied. The purpose for the testing will dictate the appropriate sampling procedures.